

ON HYPOELLIPTIC BRIDGE

XUE-MEI LI

ABSTRACT. We prove that if the Markov generator of a diffusion process satisfies the two step strong Hörmander condition, the conditioned hypoelliptic bridge satisfies an integral bound and is a continuous semi-martingale.

1. INTRODUCTION

We are motivated by the path integration formula and also by the L^2 analysis on the space of pinned continuous curves where the Brownian bridge plays an important role. Let M be a smooth connected Riemannian manifold. Denote $C([0, 1]; M)$ the space of continuous functions from $[0, 1]$ to M and $C_{x_0, z_0}([0, 1]; M)$ its subspace of curves that begin at x_0 and end at z_0 . If (x_t) is a Brownian motion with initial value x_0 and without explosion, a Brownian bridge $(b_t^{x_0, z_0}, 0 \leq t \leq 1)$ begins with x_0 and ends at z_0 is a stochastic process with probability distribution $P(\cdot | x_1 = z_0)$. The Brownian bridge is well known to induce a probability measure on $C_{x_0, z_0}([0, 1]; M)$ for M compact.

For the L^2 analysis, it is standard to equip the space with the probability measure determined by the Brownian bridge, which fuelled the study of the logarithm of the heat kernel and their derivatives. However there is no particular strong argument for the use of Brownian bridges, and indeed one is tempted to explore. For example on a Lie group, a basic object is a diffusion operator built from a family of left invariant vector fields.

If $\{X_i, i = 0, 1, \dots, m\}$ is a family of smooth vector fields, let $\mathcal{L} = \frac{1}{2} \sum_{k=1}^m L_{X_k} L_{X_k} + L_{X_0}$ where L_v denotes Lie differentiation in the direction of a vector v . If the diffusion coefficients $\{X_1, \dots, X_m\}$ and their iterated Lie brackets span the tangent space $T_x M$ at each x , \mathcal{L} is said to satisfy *the strong Hörmander condition*. Denote D_k the set of vector fields and their commutators up to level k . If \mathcal{L} satisfies the strong Hörmander condition the minimal k needed to span $T_x M$ is denoted by $l(x)$. If for all x , $l(x) \leq p$, \mathcal{L} is said to satisfy the *p -step strong Hörmander condition*. We assume that there exists a global parabolic integral kernel for \mathcal{L} , which holds if \mathcal{L} is a sub-Laplacian and the sub-Riemannian distance is complete, L. Strichartz [48]; or is symmetric and M is compact, D. Jerison and A. Sanchez-Calle [32] and B. Davies [16]; or is uniformly hypoelliptic and $M = \mathbb{R}^n$, S. Kusuoka and D. Stroock [36]. See also L. Rothschild and E. Stein [45] and G. B. Folland [24].

Suppose that \mathcal{L} satisfies the two step strong Hörmander condition and $X_0 = \sum_{k=1}^m c_k X_k$. We make the following simple observation on the hypoelliptic bridge (y_t) :

$$\mathbf{E} \int_0^1 |d \log q_{1-s}(\cdot, z_0)(X_k(y_s))| ds < \infty,$$

AMS Mathematics Subject Classification : 60Gxx, 60Hxx, 58J65, 58J70. Part of the work is done during a visit to MSRI, 2015. Address: Mathematics Institute, The University of Warwick, Coventry CV4 7AL, U.K.

in particular $(y_t, 0 \leq t \leq 1)$ is a continuous semi-martingale. The integral bound is obtained from small time estimates on the fundamental solution and its gradient, the latter from H. Cao and S.-T. Yau [13]. The Gaussian bounds for q_t depend on the volume of the intrinsic metric balls $B_x(\sqrt{t})$ for small time and on the Euclidean ball for large time. Around x the metric distance is comparable with $\rho^{\frac{1}{l(x)}}$ where ρ is the Riemannian distance. The larger is $l(x)$, the more singular is the heat kernel at 0. Hence the semi-martingale property for diffusions satisfying two-step Hörmander condition does not hint for a generalisation. It is tempting to argue that this property fails when $l(x)$ is sufficiently large. On the other hand the following results are proved recently: the Brownian bridge concentrates on the sub-Riemannian geodesic at $t \rightarrow 0$. See I. Bailleul, L. Mesnager and J. Norris [2], and Y. Inahama [31]. Since the semi-martingale property depends on properties of the heat kernel for small time, and since the sub-Riemannian geodesic is horizontal in whose direction the singularity in t should be exactly $t^{-\frac{\alpha}{2}}$, we tend to believe this semi-martingale property holds much more generally.

2. PRELIMINARIES

The purpose for this section is to familiarize ourselves with the basic properties of hypoelliptic bridges. To condition a diffusion process from x_0 to reach y_0 at 1, it is natural to assume there is a control path reaching y_0 from x_0 and the transition probability measures have positive densities, q_t , with respect to a Riemannian volume measure. Hence it is reasonable to assume Hörmander condition. It is well known that, at least when M is a compact manifold, the conditioned diffusion induces a measure on the space of continuous paths. This is noted in J. Eells and K. D. Elworthy [20], J.-M. Bismut [10], P. Malliavin and M.-P. Malliavin [40], and B. Driver [18]. Eells and Elworthy were interested in relating the Wiener and pinned Wiener measures to the topology and geometry of the path space over a manifold, which later involving the quest for an L^2 Hodge theory, see e.g. K. D. Elworthy and Xue-Mei Li [22, 23]. Bismut, Driver, Malliavin and Malliavin were interested in the quasi-invariance of the pinned Brownian motion measure. An alternative proof for the quasi-invariance theorem of Malliavin and Malliavin was given in M. Gordina [26].

We discuss two cases: in the first \mathcal{L} has an invariant measure μ , i.e. $\int \mathcal{L}f d\mu = 0$ for any $f \in C_K^\infty$, and in the second we assume estimates on the heat kernel. We begin with the first case. In general we do not know there is a global solution to $\mathcal{L}^*\mu = 0$. If \mathcal{L} satisfies strong Hörmander condition, and M is compact or \mathcal{L} is symmetric, the \mathcal{L} -diffusion (x_t) has an invariant measure. If $\{X_1, \dots, X_m\}$ are linearly independent they determine a sub-Riemannian metric. The sub-elliptic Laplacian Δ_H is defined to be $\text{trace div } \nabla^H$ where ∇^H is the sub-Riemannian gradient and the divergence is with respect to a volume form μ . Then $\Delta_H = \sum_{i=1}^m L_{X_i} L_{X_i} + X_0$ where $X_0 = -\sum_{i=1}^m \text{div}_\mu(X_i)X_i$. If the sub-Riemannian metric is complete, then Δ_H is essentially self adjoint on $L^2(M; \mu)$, see R. Strichartz [48]. If the sub-Riemannian metric is the restriction of a complete Riemannian metric or if the sub-Riemannian structure is obtained from left invariant vector fields on a lie group, M is a complete metric space with respect to the sub Riemannian distance d . In this paper we do not use a sub-Riemannian structure and will however comment on this at the end of the paper.

Throughout this paper x_t is assumed to be conservative, otherwise the set of paths considered would exclude the paths with life time less than 1, which we are not willing to compromise. For simplicity we drop the subscript 1 in q_1 . If $f : M \rightarrow \mathbb{R}$ is a differentiable function we define its horizontal gradient to be $\nabla^H f = \sum_{i=1}^m X_i df(X_i)$. Let $\hat{\mathcal{L}}$

denote the adjoint operator with respect to μ , not necessarily finite, invariant measure μ , i.e. $\int \mathcal{L}fgd\mu = \int f\hat{\mathcal{L}}gd\mu$. Denote \hat{x}_t the adjoint process.

Lemma 2.1. *If $\mathcal{L}^*\mu = 0$ has a solution and the adjoint process is conservative, the hypoelliptic bridge determines a probability measure on $C_{x_0, y_0}([0, 1]; M)$.*

Proof. Let (x_t) be an \mathcal{L} -diffusion and (y_t) the conditioned bridge process. Restricted to an interval $[0, 3/4]$, y_t is a ‘Doob transform’ of (x_t) . Let $\{w_t^i\}$ be a family of real valued independent one dimensional Brownian motions. Then x_t and y_t can be represented as solutions to the equations with initial values $x_0 = y_0$,

$$dx_t = \sum_{i=1}^m X_i(x_t) \circ dw_t^i + X_0(x_t)dt,$$

$$dy_t = \sum_{i=1}^m X_i(y_t) \circ dw_t^i + X_0(y_t)dt + \nabla^H \log q_{1-t}(y_t, y_0)dt,$$

where the gradient is with respect to the first variable.

It is easy to see that $\exp(N_t) = \frac{q_{1-t}(x_t, y_0)}{q(x_0, y_0)}$. Since $\mathbf{E} \frac{q_{1-t}(x_t, y_0)}{q(x_0, y_0)} = 1$, $\exp(N_s, 0 \leq s \leq t)$ is a martingale. If F is supported on continuous paths defined up to a time $t < 1$, then $\mathbf{E}F(y_t) = \mathbf{E}F(x_t)e^{N_t}$. From this and $\mathbf{E}N_t = 1$, we see that the finite dimensional distributions of (y_t) agree with that of the conditioned process, when restricted to $[0, t]$. Since (x_t) admits a continuous modification and hence determines a probability measure on $C([0, 3/4]; M)$, so does (y_t) .

The invariant measure π is a distributional solution to $\mathcal{L}^*m = 0$ where

$$\mathcal{L}^* = \sum_{i=1}^m L_{X_i} L_{X_i} - L_{X_0} - 2 \sum_{i=1}^m \operatorname{div}(X_i) L_{X_i} + \operatorname{div}(X_0) - \sum_{i=1}^m d(\operatorname{div}(X_i))(X_i)$$

is the L^2 adjoint of \mathcal{L} with respect to the volume measure, with respect to which the divergence is also taken. Then \mathcal{L}^* satisfies also the strong Hörmander condition. By the standard theory, see L. Hormander [29], μ has a strictly positive smooth density m .

If \hat{x}_t is adjoint to (x_t) , with respect to m , its Markov generator has the same leading term as \mathcal{L} and satisfies also strong Hörmander condition. We denote by \hat{q}_t its smooth density and there is the following identity: $m(x)q_t(x, y) = m(y)\hat{q}_t(y, x)$. Since the $\hat{\mathcal{L}}$ diffusion does not explode, we condition \hat{x}_t to reach x from y in time 1. The corresponding process is denoted by \hat{y}_t . Then \hat{y}_{1-t} has the same distribution as y_t . This follows from

$$q_{t_1, \dots, t_n}^{x_0, y_0} = \frac{q_{t_1}(x_0, x_1) \cdots q_{t_n-t_{n-1}}(x_{n-1}, x_n) q_{1-t_n}(x_n, y_0)}{q(x_0, y_0)}, \quad t_i < 1,$$

in which we replace q by \hat{q} . By the same argument as above, we see that \hat{y}_t has a continuous modification on $[0, 3/4]$. Thus x_t determines a probability measure on $C_{x_0, y_0}([0, 1]; M)$. The probability measure on the Borel σ -algebra of $M^{[0, 1]}$, agrees with those determined by the continuous modification of x_t , when restricted to paths on $[0, 3/4]$ and $[1/4, 1]$. The required conclusion follows. \square

We move on to results based on heat kernel estimates and begin with reviewing Gaussian upper bounds for the fundamental solutions. The Markov generator for an elliptic diffusion is necessarily of the form $\frac{1}{2}\Delta + Z$ where Δ is the Laplace-Beltrami operator for some Riemannian metric on M and Z is a vector field, in which case the diffusion is a Brownian motion with drift Z . Once we understand the case of $\mathcal{L} = \frac{1}{2}\Delta$, an additional (well behaved) drift vector field Z can be taken care of. For a detailed review on heat kernel upper bounds

see L. Saloff-Coste [46]. Take first $\mathcal{L} = \frac{1}{2}\Delta$. If the Ricci curvature of the manifold is bounded from below by $-K$ where K is a positive number, then $p_t(x, x) \sim t^{-\frac{n}{2}}$ where $n = \dim(M)$ and $t \in (0, 1)$. This is a theorem of P. Li and S.-T. Yau [39], extending the result of J. Cheeger and S.-T. Yau [15]. In general if there exists an increasing function $\beta : (0, \infty) \rightarrow \mathbb{R}_+$ such that for all $t > 0$ there is the on diagonal estimate $p_t(x, x) \leq \frac{1}{\beta(t)}$ and if β satisfies the doubling property, $\beta(2t) \leq A\beta(t)$ for all $t > 0$ and some number A , then for some constant D, δ , and C ,

$$p(t, x, y) \leq \frac{C}{\beta(\delta t)} e^{-\frac{\rho^2(x, y)}{2Dt}}. \quad (2.1)$$

See A. Grigoryan [27] and A. Bendikov and L. Saloff-Coste [9] for a detailed account. In the case of $M = \mathbb{R}^n$, a Sobolev inequality implies Nash's inequality which in turn implies an on diagonal estimate with $\beta(t) = t^{\frac{n}{2}}$, see J. Nash [43]. Conversely by a theorem of N. Varopoulos [49], generalised by E. Carlen, S. Kusuoka and D. Stroock [14], the on diagonal estimate implies Sobolev's inequality.

If $\mathcal{L} = \sum_{k=1}^m L_{X_k} L_{X_k} + L_{X_0}$ is not elliptic, but satisfies Hörmander condition, the bounds on the fundamental solution have different orders depending on whether the time is small or large. To use Kolmogorov's Theorem, it is for the small time we need the more refined upper bound. Under Hörmander condition the fundamental solution q_t of the parabolic equation $\frac{\partial}{\partial t} = \mathcal{L}$ is expected to admit a Gaussian upper bound. For small time, it is better to use the *intrinsic metric distance* d defined by the formula:

$$d(x, y) = \inf \left\{ l \mid \gamma : [0, l] \rightarrow M, \dot{\gamma} = \sum_{i=1}^m a_i X_i, \sum_{i=1}^m (a_i(s))^2 \leq 1 \right\}.$$

where γ is taken over all Lipschitz continuous curves on a compact interval connecting x to y . This intrinsic distance is a natural distance for \mathcal{L} , i.e. d induces the original topology of the manifold.

For diffusions on a compact manifold satisfying strong Hörmander's conditions and with the drift X_0 vanishing identically, there is the following estimates in terms of the volume of the metric ball $B_x(r\sqrt{t})$ centred at x :

$$\frac{C_1}{\text{vol}(B_x(\sqrt{t}))} e^{-\frac{C_3 d^2(x, y)}{t}} \leq q_t(x, y) \leq \frac{C_2}{\text{vol}(B_x(\sqrt{t}))} e^{-\frac{C_4 d^2(x, y)}{t}}, \quad (2.2)$$

for all $x, y \in M$ and all $t > 0$. This is a theorem of D. Jerison and A. Sanchez-Calle [32]. In A. Sanchez-Calle [47], this upper bound is obtained for (x, y) satisfying the relation $d(x, y) \leq \sqrt{t}$ and $t \leq 1$. Estimates in (2.2) for the heat kernel is effective only for small times. Indeed, as $q_t(x, y)$ is smooth and strictly positive, we obtain trivial upper and lower constant bounds for q_t . It is another matter to obtain the best constants.

For two points x, y close to each other,

$$\frac{1}{c} \rho(x, y) \leq d(x, y) \leq c \rho(x, y)^{\frac{1}{l(x)}}, \quad (2.3)$$

where $l(x)$ is the length in Hörmander's condition, assuming that the intrinsic sub-Riemannian metric associated with $\{X_1, \dots, X_m\}$ agrees with the restriction of the Riemannian metric defining ρ . If M is compact and the vector fields are BC^∞ , then d and ρ are equivalent. The upper bound for d comes from the fact that any point in a small neighbourhood of a point x , of a uniform size, can be reached from x by a controlled path. This is essentially the Box-ball theorem of A. Nagel, E. Stein S. Wainger [42]. See also R. Montgomery [41]. For symmetric diffusions on \mathbb{R}^n satisfying a 'uniform Hörmander's condition' and t small,

estimates of the above form were proved in S. Kusuoka and D. Stroock [34]. For large t the Euclidean metric is more relevant, see S. Kusuoka and D. Stroock [35]. We do not need sharp estimates on the heat kernel, however we mention that sharp estimates was obtained in E. B. Davies [16] for symmetric diffusions on a compact manifold. Also Varadhan's short time asymptotics for $\log q_t$ was given in G. Ben Arous and R. Léandre [8] and R. Léandre [38, 37]. See also P. Friz and S. De Marco [25] for a recent study.

Although an estimate of the type (2.2) is sufficient for us, the intrinsic distance is not easy to use. The fundamental solution q_t is the density of the probability distribution of the \mathcal{L} -diffusion evaluated at t with respect to the volume measure. In geodesics coordinates we easily integrate a function of ρ , not so easily a function of d . For this reason it is convenient to use the argument that established (2.3) to convert the quantities involving d^2 to ρ^2 . Let us consider the volume of the metric ball centred at x with radius \sqrt{t} . When t is sufficiently small, one could apply (2.3) for crude estimates. A much refined estimate is given in G. Ben Arous and R. Léandre [8]. For example we know that for x, y not in each other's cut locus, as $t \rightarrow 0$

$$q_t(x, y) \sim \frac{C(x, y)}{t^{\frac{n}{2}}} e^{-\frac{d^2(x, y)}{2t}}$$

On the diagonal $q_t(x, x) \sim c(x)t^{-\frac{Q(x)}{2}}$ for a number $Q(x)$ relating to $l(x)$, which holds also if X_0 is in the span of the diffusion vector fields and their first order Lie brackets. G. Ben Arous and R. Léandre gave also an example where $X_0 \neq 0$ and where q_t decreases exponentially on the diagonal.

Lemma 2.2. *Suppose that \mathcal{L} -diffusion is conservative, has a smooth density q_t and*

- (1) *For any $a_0 > 0$, $\sup_{a_0 \leq t \leq T} \sup_{x, y} q_t(x, y) < \infty$.*
- (2) *There exists positive numbers δ_0 , a and $p > 1$, s.t. for all $0 \leq s < t < T$,*

$$\begin{aligned} \sup_{s > \frac{1}{4}, |t-s| < t_0} \frac{\int_{M \times M} \rho^p(x, y) q_s(x_0, x) q_{t-s}(x, y) dy dx}{|t-s|^{1+\delta_0}} &\leq C; \\ \sup_{0 < t < \frac{3}{4}, |t-s| < t_0} \frac{\int_{M \times M} \rho^p(x, y) q_{t-s}(x, y) q_{1-t}(y, y_0) dx dy}{|t-s|^{1+\delta_0}} &\leq C. \end{aligned} \tag{2.4}$$

Then there exist positive constants t_0 and C such that for $|t-s| \leq t_0$, $\mathbf{E} \rho^p(y_s, y_t) \leq C|s-t|^{1+\delta}$.

Note we do not assume the diffusion is symmetric. By (2.1) the lemma applies to $\mathcal{L} = \frac{1}{2}\Delta$ on a complete Riemannian manifold whose Ricci curvature is bounded from below. The proof for the Lemma is included for reader's convenience.

Proof. We may assume $t_0 < 1/4$ and consider the following cases: $0 < s < t < \frac{3}{4}$; $0 < \frac{1}{4} < s < t$; $s = 0$; $t = 1$. We begin with the last case.

$$\begin{aligned} \mathbf{E} \rho^p(y_s, y_0) &= \frac{1}{q(x_0, y_0)} \int_M \rho^p(x, y_0) q_s(x_0, x) q_{1-t}(x, y_0) dx \\ &\leq \frac{\sup_{s \geq \frac{1}{4}} \sup_y q_s(x, y_0)}{q(x_0, y_0)} \int_M \rho^p(x, y_0) q_{1-s}(x, y_0) dx. \end{aligned}$$

If $0 < s < t < \frac{3}{4}$,

$$\begin{aligned} \mathbf{E}\rho^p(y_s, y_t) &= \int_M q_{1-t}(y, y_0) \int_M \frac{\rho^p(x, y) q_s(x_0, x) q_{t-s}(x, y)}{q(x_0, y_0)} dx dy \\ &\leq \frac{\sup_{t < \frac{3}{4}} \sup_y q_{1-t}(y, y_0)}{q(x_0, y_0)} \int_M \int_M q_s(x_0, x) \rho^p(x, y) q_{t-s}(x, y) dy dx, \end{aligned}$$

concluding the estimates. The estimation for the other cases are similar. To show that the finite dimensional distributions $q_t^{x_0, y_0}$ determines a probability measure on $C([0, 1]; M)$ it is sufficient to prove that there exist $p > 1$, $\delta_0 > 0$, and $t_0 > 0$ such that if $|t - s| < t_0$ and $0 \leq s \leq t \leq 1$, $\mathbf{E}\rho(y_t, y_s)^p \leq C|t - s|^{1+\delta_0}$. This completes the proof. \square

If q is a continuous and M is compact, assumption (1) is automatic. We look into condition (2) in more detail. Denote μ the Euclidean surface measure on S^n , $c_x(\xi)$ the distance to the cut point of x along the geodesic $\gamma_x(\xi)$ in the direction of $\xi \in T_x M$. Denote $ST_x M$ the unit sphere in $T_x M$ and set

$$\begin{aligned} D_x &= \{t\xi : \xi \in ST_x M, t \in [0, c(\xi))\} = T_x M \setminus C_x \\ D_x(r) &= \{\xi \in ST_x M : r < c(\xi)\}. \end{aligned}$$

where C_x is the Riemannian cut locus at x . Note that $D_x(r)$ decrease with r . On D_x , \exp_x is a diffeomorphism onto its image. Denote $J_x(v)$ the determinant of $(d\exp_x)_v$ identifying the tangent spaces of $T_x M$ with itself. Furthermore we denote $A_x(r)$ the lower area function:

$$A(x, r) = \int_{D_x(r)} J_x(r\xi) d\mu(\xi) = \frac{1}{r^{n-1}} \int_{D_x} J_x(\eta) d\mu(\eta).$$

If $A(y_0, r)$ is bounded then the last inequality in the Lemma below holds trivially.

Lemma 2.3. *Suppose that there exist positive constants $C_1, C_2, C_3, \alpha, a, t_0 < 1$, positive increasing real valued functions β_i decaying at most polynomially near 0, such that the following estimates hold for $t < t_0$,*

$$\begin{aligned} q_t(x, y) &\leq \frac{C_1}{\beta_2(t)}, \quad q_t(x, y) \leq \frac{C_1}{\beta_1(t)} e^{-\frac{C_2 \rho^{2\alpha}(x, y)}{t}} \text{ when } \rho(x, y) \geq a\sqrt{t}; \\ \sup_{u \geq 0} \int_{au}^{\infty} r^{\frac{p+n}{\alpha}} e^{-C_2 r^2} A(x, r^{\frac{1}{\alpha}} u^{\frac{1}{\alpha}}) dr &< \infty. \end{aligned}$$

Then assumption (2) of Proposition 2.2 holds.

Proof. Let us consider $p > 1$, $0 \leq s \leq t \leq \frac{3}{4}$ and $|t - s| \leq t_0$. The other cases are similar. Working in polar coordinates we see that

$$\begin{aligned} &\int_M q_s(x_0, x) \int_M \rho^p(x, y) q_{t-s}(x, y) dy dx \\ &= \int_M q_s(x_0, x) \int_0^\infty r^p \int_{D_x(r)} q_{t-s}(y, \exp_x(r\xi)) J_x(r\xi) \mu(d\xi) r^{n-1} dr dx. \end{aligned}$$

We plug in the assumed upper bounds for the heat kernel in the respective regions to see the right hand side is bounded by:

$$\begin{aligned} &\int_M q_s(x_0, x) \int_0^{a\sqrt{t-s}} r^{n+p-1} \frac{C_1}{\beta_2(t-s)} \int_{D_x(r)} J_x(r\xi) \mu(d\xi) dr dx \\ &+ \int_M q_s(x_0, x) \frac{C_1}{\beta_1(t-s)} \int_{a\sqrt{t-s}}^\infty r^{n+p-1} e^{-\frac{C_2 r^{2\alpha}}{t-s}} \int_{D_x(r)} J_x(r\xi) \mu(d\xi) dr dx, \end{aligned}$$

which is further bounded by

$$\begin{aligned} & \frac{C_1}{\beta_2(t-s)} a^{n+p-1} (t-s)^{\frac{n+p-1}{2}} \int_M q_s(x_0, x) dx \int_0^{a\sqrt{t-s}} A(x, r) dr \\ & + \frac{C_1}{\beta_1(t-s)} \int_M dx q_s(x_0, x) \int_{a\sqrt{t-s}}^\infty r^{p+n-1} e^{-\frac{C_2 r^{2\alpha}}{t-s}} A(x, r) dr. \end{aligned}$$

This means,

$$\begin{aligned} & \int_M q_s(x_0, x) \int_M \rho^p(x, y) q_{t-s}(x, y) dy dx \\ & \leq \frac{C_1 a^{n+p-1} (t-s)^{\frac{n+p-1}{2}}}{\beta_2(t-s)} \int_M q_s(x_0, x) \int_0^{a\sqrt{t-s}} A(x, r) dr dx \\ & + \frac{C_1 (t-s)^{\frac{p+n}{2\alpha}}}{\beta_1(t-s)} \int_M dx q_s(x_0, x) \int_{a\sqrt{t-s}}^\infty r^{\frac{p+n}{\alpha}} e^{-C_2 r^2} A(x, r^{\frac{1}{\alpha}} (t-s)^{\frac{1}{2\alpha}}) dr. \end{aligned}$$

Since $\beta_1(t), \beta_2(t)$ decays at most polynomially near 0, we may choose p and $\delta > 0$ such that the assumption (2) of Proposition 2.2 holds. \square

The conclusions of Lemma 2.2 and 2.3 hold if $M = \mathbb{R}^n$, \mathcal{L} satisfies the following Kusuoka-Stroock's uniform Hörmander's condition: there exists an integer p such that $l(x) \leq p$. The vector fields $\{X_1, \dots, X_m\}$ and their iterated brackets up to order p give rise to a $n \times n$ symmetric matrix that is uniformly elliptic on \mathbb{R}^n . Also X_0 is in the linear span of $\{X_1, \dots, X_m\}$. In fact, there exist constants $M > 1$ and r_0 such that for any $t \in (0, 1]$ and $x, y \in \mathbb{R}^n$, [34], the upper bound in (Gaussian-bounds) holds with $C_4 = \frac{1}{C_2}$. Also the lower surface function $A(x, r)$ is bounded by a constant, the last inequality in Lemma 2.3 is satisfied. Assumption (2) in Lemma 2.2 holds. For $t \geq 1$, S. Kusuoka and D. Stroock proved the following [35], $q_t(x, y) \leq M t^{-\frac{n}{2}} e^{-\frac{|y-x|^2}{Mt}}$, which ensures assumption (1) in Lemma 2.2.

3. THE SEMI-MARTINGALE PROPERTY

Let $x_0, z_0 \in M$ and $(y_t, 0 \leq t < 1)$ be the solution of the following equation

$$dy_t = \sum_{i=1}^m X_i(y_t) \circ dw_t^i + X_0(y_t) dt + \nabla^H \log q_{1-t}(\cdot, z_0)(y_t) dt, \quad y_0(\omega) = x_0$$

Theorem 3.1. *If M is compact, X_0 is divergence free, and \mathcal{L} satisfies the two step strong Hörmander condition, then for each $i = 1, \dots, m$,*

$$\mathbf{E} \int_0^1 |d \log q_{1-s}(\cdot, z_0)(X_i(y_s))| ds < \infty.$$

If $\mathcal{L} = \frac{1}{2} \Delta$, this is well known. The standard proof relies on the following estimate on the heat kernel: $|\nabla_x \log p_t(x, y)| \leq C(\frac{1}{\sqrt{t}} + \frac{\rho(x, y)}{t})$, which can be proved probabilistically or follows from the Gaussian type upper and lower bounds and Hamilton's estimate for the heat kernel, R. Hamilton[28]:

$$s |\nabla_x \log p_s(\cdot, y)|^2 \leq C_1 \log\left(\frac{C_2}{s^{\frac{n}{2}} p_s(\cdot, y_0)}\right).$$

See B. Driver [18] and the following books and survey: Bismut [11], B. Driver [19] and E. Hsu [30]. In B. Kim [33, Prop. 5.2] the following inequality is proved for a positive

bounded smooth solution, satisfying further suitable L^2 estimates: $t|\nabla \ln u(x, t)| \leq C(1 + t) \ln(\frac{M}{u(x, t)})$. There \mathcal{L} is a ‘sub-elliptic’ operator. If we apply this to the kernel q_t , together with a favourable Gaussian lower bound for $\nabla \ln u$, e.g. (2.2), assuming that the metric balls of volume t is polynomial in t , we have

$$|\nabla_x \log q_t(x, y)|^2 \leq C\left(\frac{|\ln t|}{t} + \frac{\rho^2(x, y)}{t^2}\right).$$

In terms of integrability this estimate is slightly better than the corresponding one in [33, Lemma 5.3]. However it is still on the wrong side of critical integrability at $t = 0$.

We give some examples where the theorem holds. (1) $M = SU(2)$, and X_1^*, X_2^* are left invariant vector fields generated by two Pauli matrices. (2) M is the torus, $X_1(x, y) = \frac{\partial}{\partial x}$ and $X_2(x, y) = \sin(2\pi x) \frac{\partial}{\partial y}$. (3) $M = G/Z^3$ where G is the Heisenberg group and $X_1(x, y, z) = \frac{\partial}{\partial x}$ and $X_2(x, y, z) = \frac{\partial}{\partial y} + x \frac{\partial}{\partial z}$.

Proof. It is sufficient to prove that $\int_0^1 \sqrt{\mathbf{E}|\nabla \log q_{1-s}(y_s, z_0)|^2} ds < \infty$. We use the following theorem of H. Cao and S. T. Yau [13]. Let X_0, X_1, \dots, X_m be smooth vector fields on a compact manifold such that $X_0 = \sum_{k=1}^m c_k X_k$ for a set of smooth real valued functions c_k on M . Likewise suppose that for every set of $i, j, k = 1, \dots, m$, $[[X_i, X_j], X_k](x)$ can be expressed as a linear combination of vector fields from $\{X_{i'}, [X_{j'}, X_{k'}], i', j', k' = 1, \dots, m\}$. If u_t is a positive solution to the equation $\frac{\partial}{\partial t} u_t = \sum_i L_{X_i} L_{X_i} + L_{X_0}$, there exists a constant $\delta_0 > 1$, such that for all $\delta > \delta_0$ and $t > 0$,

$$\frac{1}{u^2} \sum_i |L_{X_i} u|^2 \leq \delta \frac{L_{X_0} u}{u} + \delta \frac{1}{u} \frac{\partial u}{\partial t} + \frac{C_1}{t} + C_2,$$

where C_1, C_2 are constants depending on \mathcal{L} and δ_0 . Applying this to the fundamental solution q_t , we see that

$$\mathbf{E}|\nabla \log q_{1-s}(y_s, z_0)|^2 \leq \delta \mathbf{E} \frac{L_{X_0} q_{1-s}(\cdot, z_0)}{q_{1-s}(\cdot, z_0)}(y_s) + \delta \mathbf{E} \frac{\frac{\partial q_{1-s}(\cdot, z_0)}{\partial s}(y_s)}{q_{1-s}(y_s, z_0)} + \frac{C_1}{1-s} + C_2.$$

Using the explicit formula for the probability density of y_t , we see that for any $s < 1$,

$$\begin{aligned} \mathbf{E} \left(\frac{\frac{\partial}{\partial s} q_{1-s}(\cdot, z_0)(y_s)}{q_{1-s}(y_s, z_0)} \right) &= \int_M \frac{\frac{\partial}{\partial s} q_{1-s}(x, z_0) q_s(x_0, x)}{q(x_0, z_0)} dx \\ &= \int_M \frac{\frac{\partial}{\partial s} (q_{1-s}(x, z_0) q_s(x_0, x)) - q_{1-s}(x, z_0) \frac{\partial}{\partial s} q_s(x_0, x)}{q(x_0, z_0)} dx = - \int_M \frac{q_{1-s}(x, z_0) \frac{\partial}{\partial s} q_s(x_0, x)}{q(x_0, z_0)} dx. \end{aligned}$$

Since the divergence of X_0 vanishes, the same reasoning leads to the following identities:

$$\begin{aligned} \mathbf{E} \left(\frac{L_{X_0} q_{1-s}(\cdot, z_0)}{q_{1-s}(\cdot, z_0)}(y_s) \right) &= \int_M \frac{L_{X_0} q_{1-s}(x, z_0) q_s(x_0, x)}{q(x_0, z_0)} dx \\ &= \int_M \frac{L_{X_0} (q_{1-s}(x, z_0) q_s(x_0, x)) - q_{1-s}(x, z_0) L_{X_0} q_s(x_0, x)}{q(x_0, z_0)} dx = \int_M \frac{-q_{1-s}(x, z_0) L_{X_0} q_s(x_0, x)}{q(x_0, z_0)} dx \end{aligned}$$

Let us consider the integral from $\frac{1}{2}$ to 1.

$$\begin{aligned} &\int_{\frac{1}{2}}^1 \sqrt{\mathbf{E}|\nabla \log q_{1-s}(y_s, z_0)|^2} ds \\ &\leq \int_{\frac{1}{2}}^1 \left(\int_M \left| \frac{q_{1-s}(x, z_0) (L_{X_0} q_s(x_0, x) + \frac{\partial}{\partial s} q_s(x_0, x))}{q(x_0, z_0)} \right| dx + \frac{C_1}{1-s} + C_2 \right)^{\frac{1}{2}} ds \end{aligned}$$

Since q_t is smooth and the manifold is compact, there is a constant C_3 such that

$$\sup_{s \in [\frac{1}{2}, 1]} \left| L_{X_0} q_s(x_0, x) + \frac{\partial}{\partial s} q_s(x_0, x) \right| \leq C_3,$$

$$\int_{\frac{1}{2}}^1 \sqrt{\mathbf{E} |\nabla \log q_{1-s}(y_s, z_0)|^2} ds \leq \int_{\frac{1}{2}}^1 \sqrt{\frac{C_3}{q(x_0, z_0)} + \frac{C_1}{1-s} + C_2} ds < \infty.$$

The same reasoning shows that $\int_0^{\frac{1}{2}} \sqrt{\mathbf{E} |\nabla \log q_{1-s}(y_s, z_0)|^2} ds$ is finite. \square

Remark 3.2. (1) If $\mathcal{L} = \frac{1}{2}\Delta$, and M is a complete Riemannian manifold with Ricci curvature is non-negative, there is the Harnack inequality: $\frac{|\nabla u|^2}{u^2} - \alpha \frac{u_t}{u} \leq \alpha^2 \frac{n}{2t}$ where $\alpha > 1$ and C are constants. See P. Li and S.-T. Yau [39] and B. Davies [17]. Hence the proof of the theorem applies. (2) Two step Hormander condition is used in J. Picard [44], for a different problem. (3) It is also interesting to explore the Cameron-Martin quasi-invariance theorem in this context and prove the flow of the SDE is quasi invariant under a Girsanov-Martin shift. This should be straightforward if the shift is induced from vector fields of the form $\int_0^1 X^i(x) h_s^i ds$. The quasi-invariance of the conditioned hypoelliptic measure is now known in some sub-Riemannian case, see F. Baudoin, M. Gordina and M. Tai [6] for Heisenberg type Lie groups. (4) Finally we remark that a limited Li-Yau type inequality in F. Baudoin and N. Garofalo [5], see also F. Baudoin, M. Bonnefont and N. Garofalo [4], was extended to certain sub-Riemannian situation, we have not yet managed to use it to our advantage, and this will be for a future study. A study for semigroups of Hörmander type second order differential operators, not necessarily satisfying Hörmander condition, can be found in K. D. Elworthy, Y. LeJan and Xue-Mei Li [21]. Finally we refer to the following articles and book for further analysis on and in sub-Riemannian geometry: A. Agrachev and D. Barilari [1], N. Varopoulos, L. Saloff-Coste, and T. Coulhon [50], M. Bramanti [12], A. Bellaïche [7], D. Barilari, U. Boscain and R. Neel [3] and the book by R. Montgomery [41].

REFERENCES

- [1] A. Agrachev and D. Barilari. Sub-Riemannian structures on 3D Lie groups. *J. Dyn. Control Syst.*, 18(1):21–44, 2012.
- [2] I. Bailleul, L. Mesnager, and J. Norris. Small-time fluctuations for the bridge of a sub-riemannian diffusion. arXiv:1505.03464, 2015.
- [3] Davide Barilari, Ugo Boscain, and Robert W. Neel. Small-time heat kernel asymptotics at the sub-Riemannian cut locus. *J. Differential Geom.*, 92(3):373–416, 2012.
- [4] Fabrice Baudoin, Michel Bonnefont, and Nicola Garofalo. A sub-Riemannian curvature-dimension inequality, volume doubling property and the Poincaré inequality. *Math. Ann.*, 358(3-4):833–860, 2014.
- [5] Fabrice Baudoin and Nicola Garofalo. Curvature-dimension inequalities and ricci lower bounds for sub-riemannian manifolds with transverse symmetries. arXiv:1101.3590, 2014.
- [6] Fabrice Baudoin, Maria Gordina, and Tai Melcher. Quasi-invariance for heat kernel measures on sub-Riemannian infinite-dimensional Heisenberg groups. *Trans. Amer. Math. Soc.*, 365(8):4313–4350, 2013.
- [7] A. Bellaïche. The tangent space in sub-Riemannian geometry. *J. Math. Sci. (New York)*, 83(4):461–476, 1997. Dynamical systems, 3.
- [8] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Probab. Theory Related Fields*, 90(3):377–402, 1991.
- [9] A. Bendikov and L. Saloff-Coste. On- and off-diagonal heat kernel behaviors on certain infinite dimensional local Dirichlet spaces. *Amer. J. Math.*, 122(6):1205–1263, 2000.
- [10] Jean-Michel Bismut. Martingales, the Malliavin calculus and hypoellipticity under general Hörmander’s conditions. *Z. Wahrsch. Verw. Gebiete*, 56(4):469–505, 1981.
- [11] Jean-Michel Bismut. *Large deviations and the Malliavin calculus*, volume 45 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1984.

- [12] Marco Bramanti. *An invitation to hypoelliptic operators and Hörmander's vector fields*. Springer Briefs in Mathematics. Springer, Cham, 2014.
- [13] Huai Dong Cao and Shing-Tung Yau. Gradient estimates, Harnack inequalities and estimates for heat kernels of the sum of squares of vector fields. *Math. Z.*, 211(3):485–504, 1992.
- [14] E. A. Carlen, S. Kusuoka, and D. W. Stroock. Upper bounds for symmetric Markov transition functions. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(2, suppl.):245–287, 1987.
- [15] Jeff Cheeger and Shing Tung Yau. A lower bound for the heat kernel. *Comm. Pure Appl. Math.*, 34(4):465–480, 1981.
- [16] E. B. Davies. Gaussian upper bounds for the heat kernels of some second-order operators on Riemannian manifolds. *J. Funct. Anal.*, 80(1):16–32, 1988.
- [17] E. B. Davies. *Heat kernels and spectral theory*, volume 92 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1989.
- [18] Bruce K. Driver. A Cameron-Martin type quasi-invariance theorem for Brownian motion on a compact Riemannian manifold. *J. Funct. Anal.*, 110(2):272–376, 1992.
- [19] Bruce K. Driver. Curved Wiener space analysis. In *Real and stochastic analysis*, Trends Math., pages 43–198. Birkhäuser Boston, Boston, MA, 2004.
- [20] J. Eells and K. D. Elworthy. Wiener integration on certain manifolds. In *Problems in non-linear analysis (C.I.M.E., IV Ciclo, Varenna, 1970)*, pages 67–94. Edizioni Cremonese, Rome, 1971.
- [21] K. D. Elworthy, Y. Le Jan, and Xue-Mei Li. *On the geometry of diffusion operators and stochastic flows*, volume 1720 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999.
- [22] K. D. Elworthy and Xue-Mei Li. An L^2 theory for differential forms on path spaces. I. *J. Funct. Anal.*, 254(1):196–245, 2008.
- [23] K. David Elworthy and Xue-Mei Li. Geometric stochastic analysis on path spaces. In *International Congress of Mathematicians. Vol. III*, pages 575–594. Eur. Math. Soc., Zürich, 2006.
- [24] G. B. Folland. Subelliptic estimates and function spaces on nilpotent Lie groups. *Ark. Mat.*, 13(2):161–207, 1975.
- [25] P. Friz and S. De Marco. Varadhan's estimates, projected diffusions, and local volatilities. arXiv:1311.1545, 2013.
- [26] Maria Gordina. Quasi-invariance for the pinned Brownian motion on a Lie group. *Stochastic Process. Appl.*, 104(2):243–257, 2003.
- [27] Alexander Grigor'yan. *Heat kernel and analysis on manifolds*, volume 47 of *AMS/IP Studies in Advanced Mathematics*. American Mathematical Society, Providence, RI; International Press, Boston, MA, 2009.
- [28] Richard S. Hamilton. A matrix Harnack estimate for the heat equation. *Comm. Anal. Geom.*, 1(1):113–126, 1993.
- [29] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.
- [30] Elton P. Hsu. *Stochastic analysis on manifolds*, volume 38 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
- [31] Y. Inahama. Large deviations for rough path lifts of Watanabe's pullbacks of delta functions. arXiv:1412.8113, 2014.
- [32] David S. Jerison and Antonio Sánchez-Calle. Estimates for the heat kernel for a sum of squares of vector fields. *Indiana Univ. Math. J.*, 35(4):835–854, 1986.
- [33] B. Kim. Poincaré inequality and the uniqueness of solutions for the heat equation associated with subelliptic diffusion operators. arXiv:1305.0508, 2013.
- [34] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 34(2):391–442, 1987.
- [35] S. Kusuoka and D. Stroock. Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator. *Ann. of Math. (2)*, 127(1):165–189, 1988.
- [36] Shigeo Kusuoka and Daniel Stroock. Applications of the Malliavin calculus. I. In *Stochastic analysis (Katata/Kyoto, 1982)*, volume 32 of *North-Holland Math. Library*, pages 271–306. North-Holland, Amsterdam, 1984.
- [37] R. Léandre. Uniform upper bounds for hypoelliptic kernels with drift. *J. Math. Kyoto Univ.*, 34(2):263–271, 1994.
- [38] Rémi Léandre. Majoration en temps petit de la densité d'une diffusion dégénérée. *Probab. Theory Related Fields*, 74(2):289–294, 1987.
- [39] Peter Li and Shing-Tung Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156(3-4):153–201, 1986.
- [40] Marie-Paule Malliavin and Paul Malliavin. Integration on loop groups. I. Quasi invariant measures. *J. Funct. Anal.*, 93(1):207–237, 1990.

- [41] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.
- [42] Alexander Nagel, Elias M. Stein, and Stephen Wainger. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.*, 155(1-2):103–147, 1985.
- [43] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.
- [44] Jean Picard. Gradient estimates for some diffusion semigroups. *Probab. Theory Related Fields*, 122(4):593–612, 2002.
- [45] Linda Preiss Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.
- [46] Laurent Saloff-Coste. The heat kernel and its estimates. In *Probabilistic approach to geometry*, volume 57 of *Adv. Stud. Pure Math.*, pages 405–436. Math. Soc. Japan, Tokyo, 2010.
- [47] Antonio Sánchez-Calle. Fundamental solutions and geometry of the sum of squares of vector fields. *Invent. Math.*, 78(1):143–160, 1984.
- [48] Robert S. Strichartz. Sub-Riemannian geometry. *J. Differential Geom.*, 24(2):221–263, 1986.
- [49] N. Th. Varopoulos. Hardy-Littlewood theory for semigroups. *J. Funct. Anal.*, 63(2):240–260, 1985.
- [50] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon. *Analysis and geometry on groups*, volume 100 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1992.